

人脸动画方法综述*

潘红艳^{1,2}, 柳杨华¹, 徐光祐¹

(1. 清华大学 计算机科学与技术系, 北京 100084; 2. 天津广播电视大学 计算机教研室, 天津 300191)

摘要: 在分析人脸动画实现的困难点的基础上, 介绍了人脸动画的三大类实现方法: 基于生理模型、基于三维模型和基于图像的方法。探讨了这些方法的特点, 并指出了它们各自的适用领域和局限性, 提出了各类人脸动画方法需要根据实际需求相互灵活结合使用的观点。

关键词: 人脸动画; 基于生理模型; 基于三维模型; 基于图像

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1001-3695(2008)02-0327-05

Review on methods of facial synthesis

PAN Hong-yan^{1,2}, LU Yang-hua¹, XU Guang-you¹

(1. Dept of Computer Science & Technology, Tsinghua University, Beijing 100084, China; 2. Teaching Center, Tianjin Broadcast & TV University, Tianjin 300191, China)

Abstract: This paper analyzed difficulties of realization of facial synthesis and introduced three kind of realistic methods about facial synthesis: method based on biometric model, method based on 3D facial synthesis model and method based on image. Discussed characteristic of the methods, then pointed out applicable field and localization of each method. Put forward the point of view that all kinds of methods of facial synthesis should combined with actual command.

Key words: facial synthesis; based on biometric model; based on 3D facial synthesis model; based on image

人脸动画是指通过计算机合成的人脸表情运动。早在 1972 年, Parke 首创了人脸动画系统。通过测绘技术建立网格模型, 对不同的表情进行插值, 向人们初步呈现了人脸动画的魅力。在人际交流中, 人脸是人类情感表达和交流最重要、最直接的载体之一。1989 年, 心理学家的研究表明, 物理世界中的人与人之间的直接交流主要有文本传递、语音传递 (听觉) 和表情体态传递 (视觉) 三种模式。由此, 心理学家提出一个公式: 情感表达 = 7% 文本 + 39% 语音 + 54% 表情动作^[1]。可见, 人的交流不仅是多模态的, 而且非语言模式占有重要地位。

人脸动画的应用领域和前景非常广泛。Pandzic 等人^[2]通过对 190 人的实验调查证明, 人脸配合运动能够帮助用户理解噪声环境下的讲话内容, 使用户更加集中注意力, 使服务对用户更具有吸引力, 并且能够有助于用户接受更长的等待时间。让用户面对提供人脸动画的服务以及面对不提供人脸动画的相同服务, 对比的效果相当明显。动画的人脸能够给计算机带来人性化和亲和力, 给计算机提供的交互界面营造友好和谐的氛围, 并且在聋哑人辅助教学系统中的作用尤其重要^[3]。人脸动画技术还具有很大的商业潜力, 对于电子商务、顾客关系维护管理 (customer relation management, CRM) 等都具有很大的促进作用, 增强商家与用户之间的沟通与理解。同时人脸动画还被应用于虚拟现实、虚拟主持人、电视制作、游戏娱乐等很多领域。尽管目前人脸动画的表情自动生成技术还受到各种限制, 但已经有一些公司开始将其产业化。中央电视台第六套电影频道节目中也采用虚拟主持人来给观众介绍节目内容; RedTed 公司开发出一种三维人脸模型, 可用文本、语音甚至音

乐进行控制; Sylvie 公司推出了 Verbot (verbally enhanced software robot), 这一人脸模型能够实现三维空间的智能人物与人的交流。

人脸动画越来越多地应用于面向以计算机为媒介的人际交流 (computer mediated communication, CMC)。其作用是扮演代表远方合作者的视觉窗口或替身的角色, 以方便人们通过网络进行远程交流沟通。特别是使人们能够在低网络带宽或在很低的传输码流条件下, 通过计算设备 (包括台式机、笔记本以及手机、PDA 等各种移动设备) 在人与人之间方便地进行交流和合作。其中包括虚拟视频会议和可视电话、远程教学等。在视频会议和可视电话中的人脸动画技术, 使得人们可以在低带宽网络条件下也能够远程地进行相对流畅的面对面交流, 拉近了人与人之间的距离。通过人脸动画, 人们一方面可以通过电脑替身表达自己的思想和情绪, 及时提供和反馈信息, 使他人感觉到有意义的沟通过程; 另一方面又能够根据自己的意愿隐藏自己当前交流的真实所处环境, 保护个人隐私。

人脸动画的难点

人脸动画研究领域的主要研究内容包括对人脸的几何形状与纹理表现模式进行建模, 并建立人脸表情、动作等动态信息的描述和表示方法, 以求生成能够满足各种需求的人脸表情和运动。

目前, 人脸动画还没有得到广泛运用和推广的原因是由于人脸的特殊性造成了人脸动画的困难。首先, 从人脸几何结构来看, 它是复杂的三维曲面, 其几何形状非常复杂, 并且每个人

收稿日期: 2006-10-24; 修回日期: 2007-03-13 基金项目: 国家自然科学基金资助项目 (60273005)

作者简介: 潘红艳 (1967-), 女, 天津人, 副教授, 主要研究方向为普适计算、软件工程 (panhongyan0001@126.com); 柳杨华 (1979-), 女, 博士, 主要研究方向为计算机视觉等; 徐光祐 (1941-), 男, 上海人, 博导, 主要研究方向为计算机视觉、普适计算等。

的脸部形状和纹理各不相同。人脸的各个器官,如眼睛、鼻子、嘴巴等的形状变化非常大,很难用模型精确描述;并且人脸的表面不但具有无数细小的皱纹,还会呈现微妙的颜色和纹理的变化。其次,从人脸的生理结构和运动机理机制上来看,人脸是一个层次化结构,它由头骨、肌肉层、覆盖的结缔组织和外部皮肤层组成。人脸表情的产生是由脸部多个肌肉群的协作变形引起的,涉及人脸面部多个器官的协调运动。产生人脸表情或动作涉及到的主要肌肉有口轮匝肌、鼻肌、颧肌、眼轮匝肌、皱眉肌等。例如唇部动作主要由口轮匝肌来完成;皱鼻子等动作主要是由鼻肌的收缩产生;颧肌的收缩会造成颧骨处脸颊的提高;眼睛及眼部周围的变化主要是由眼轮匝肌收缩导致的。如此复杂的运动机理很难用精确的计算机语言描述^[4]。再者,人脸的特殊材质决定了它的特殊光照特性。皮下的不规则散射、折射以及脸部毫毛对光照的影响使光照情况更加复杂化。最与众不同的是人对人脸生来就具有特别的敏感程度。尤其是对于周围熟悉的人,任何微妙的表情变化都能够觉察出来。这就使得人脸建模与动画变得更加困难。

按照传统的 Mar 视觉计算理论,人脸动画需要根据影调、运动、立体、轮廓、纹理等线索恢复人脸表面的形状,并且对人脸建立精确的三维模型表示,这是相当困难的事情。因此不少研究人员尝试根据需求应用,对人脸的建模和动画作面向任务的简化。这是实现具有实际应用价值人脸动画的有活力的思路。

从难易程度上对人脸动画方法分类

根据实现动画难易程度,人脸动画由易到难可以分为三个层次,如图 1 所示。

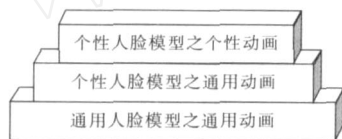


图 1 人脸动画的三个实现层次

1)通用人脸模型之通用动画 对于所有的动画人脸采用相同的模型,并且对于动画中动作表情变化的刻画也采用相同的模式。

2)个性人脸模型之通用动画 对不同的动画对象建立不同的动画模型,但是表情动作的运动模式是相同的。

3)个性人脸模型之个性动画 不但对于不同的动画对象有独特的动画模型,其动画的模式也是根据每个对象不同的特点建立的。每个层次上实现人脸动画所需关注的关键问题和运用的关键技术都不尽相同。

从实现上对人脸动画方法分类

从人脸动画的实现方法的角度来看,现有的人脸动画方法可以大致分为基于生理模型、基于三维模型(3D model-based)和基于图像(image-based)的三大类人脸动画方法。

基于生理模型的人脸动画方法

人脸是骨头、软骨、肌肉、神经、血管、腺体、脂肪组织、结缔组织、皮肤和头发等组成的复杂集合。按照生物解剖学理论,可以将人脸组织简化成皮肤、肌肉和骨骼三个层次。人脸的表

情运动主要通过肌肉层的弹性作用控制,通过皮肤层来表现。基于生理模型的人脸动画通过人脸的解剖机理和物理结构,模拟人脸的肌肉构成和运动构造简化的人脸模型,以生成脸部的表情动作。这种人脸建模动画的思想非常自然。

Platt和 Badler^[3]认为人脸可以用弹性网格结构表示,脸部运动可以通过肌肉将力传递作用于脸部的弹性网格结构来模拟。他们采用 38 条肌肉块作用于表示人脸的弹性网格结构使之产生变形,生成人脸动作。1987 年,Waters^[5]通过肌肉向量域的方向和原点,以及肌肉向量插入点的位置来定义人脸肌肉,用余弦函数来描述肌肉向量的域,将其嵌入在三维网格模型的特定位置,构造圆锥形类似于高度场的空间。Magenat-Thalman 等人^[6]则采用抽象肌肉动作过程(abstract muscle action,AMA),通过对抽象肌肉的控制模拟脸部表情。之后,Terzopoulos 与 Waters^[7]合作提出了更为细致的肌肉模型——多层弹性网格肌肉模型。Lee 等人在 Waters 工作的基础上提出了基于物理皮肤和肌肉的综合模型^[4,8]。

基于生理模型的人脸动画,通过引入肌肉模型以及驱动人脸行为的内在变形机制,使模拟的人脸真实感得到了一定的保障。但是由于对人脸生理结构还在深入研究当中,基于生理的人脸模型暂时时还不可能达到精准的水平,很难定量地给出人脸的准确构造和运动方式。这些方法所采用的人脸模型复杂,使得对不同的人脸部建模相当困难,动画生成中不可避免地涉及到巨大的计算量,并且每个人脸部的独特性给这些方法带来了更大的阻扰。因为不但每个人的肌肉粗壮程度和肌肉力量不同,而且动作习惯大不相同会导致肌肉运动习惯也大不一样。这些问题都很难用简单的肌肉模型去描述。

尽管后来又有不少研究人员提出类肌肉模型的方法。这些方法不涉及复杂的人脸生理结构,而采用类似肌肉的方式对人脸网格进行变形,以模仿人脸肌肉的运动,如伪肌肉模型^[6]、基于样条的模型^[9]和有限元模型^[10]。虽然它们相对于采用肌肉模型的人脸动画方法而言,在某种程度上简化了人脸模型,也降低了计算负担,但是由于它们缺乏对人脸纹理的建模,不能生成逼真的动画人脸纹理;并且由于对人脸的运动缺乏控制能力,使得人脸表情的生成不易调节。

基于三维模型的人脸动画方法

基于三维模型的人脸动画方法,以三维人脸模型作为动画基础来实现人脸动画。通过三维人脸模型对脸部的外观和动作模式进行建模,利用不同的函数或者参数的变化控制合成不同的人脸动作和表情。三维人脸模型的设计选取决定了人脸动画的效果、实现的难易程度以及动画效率。基于三维模型的人脸动画一般分为建模和合成两个阶段。在建模阶段,根据已知的人脸结构、形状等先验知识以及建立三维模型所需要的各种条件、参数、数据等要求,对输入的图像或是图像序列进行图像分析和处理,如边界检测、姿态估计、人脸特征抽取、特征检测、特征跟踪等,以获得相应的模型参数。合成阶段是在一定的动画规则的基础上,根据所需要的动画要求给出控制三维模型动画所需要的函数表达或形状纹理参数,以驱动模型获得动画图像。因此基于三维模型的人脸动画方法要解决两个方面的问题:三维人脸建模和动画驱动。

三维人脸建模方法

建立人脸的三维模型需要获取稠密的人脸三维信息,包括

人脸的几何信息和纹理信息。这些三维信息可以利用复杂精细的设备,只经过一般的配准和立体视觉求解获取^[11~15];也可以通过普通的图像获取设备采集图像或图像序列,基于计算机视觉(立体视觉、影调、边缘检测等方法)和图形图像学算法通过复杂的运算从图像中提取立体信息^[16~19]。

稠密的三维数据可以由专门的三维数据采集设备获取。Lee等人^[11]通过三维激光扫描仪 Cyberware 得到人脸稠密的深度图和纹理图;Hong等人^[12]和 Jeong等人^[13]也采用类似的方法构造人脸模型。此方法能够获得高精度的三维信息,但是由于三维激光扫描仪过于昂贵,很少应用于平常的工作和生活中。相类似的还有基于结构光的主动视觉方法来获取三维信息^[14,15],借助人光源在一定程度上降低了立体视觉中对应点匹配的难度。过于复杂的设备或者过多的人工干涉成为了这些方法被广泛应用的阻碍。

为了能够在通常条件下利用普通的图像捕捉设备进行三维数据的提取,许多研究人员基于图像摄影测量技术,根据计算机视觉中影调、运动、立体、轮廓、纹理等线索恢复物体表面的形状,通过对图像或者图像序列的分析获取稠密的三维信息。稠密的三维数据可以通过多部摄像机的二维图像的稠密对应,根据多目立体视觉原理求解^[16,17]。Shan等人^[18]采用了类似的方法。这类方法采用的图像捕捉设备简单且较少有人工介入,从方便用户使用的角度而言很有实用价值。但是由于基于角点检测的特征点匹配,在脸部的皮肤光滑区域容易产生误差;而且通过多视图的多组帧间对应去恢复三维的纹理,这种优化问题对初始值尤为敏感且误差容易积累。

Vetter等人^[19]提出单幅人脸图像的方法。由于根据光流在人脸图像之间作稠密的匹配对应,极易受到采集图像环境中光照和噪声的影响。Pighin^[20]采用同时多视角(5个视角)采集多幅人脸图像的方法,在这些人脸图像上分别标定少量人脸特征点(在正面图像上选择 13 个人脸特征点),通过这组图像对应特征点去估计摄像机内外参数以及特征点在空间的三维位置。在对已有特征点进行插值得到更多的顶点位置之后,通过增加新的形状对应约束,对通用人脸模型作逐步的细化调整,得到特定人的脸部三维模型(图 2)。这种方法能够获得较为准确的三维人脸模型,但同时从多视角获取人脸图像需要多部照相设备的配合,并且需要较多的人工参与标定出准确的人脸特征点。

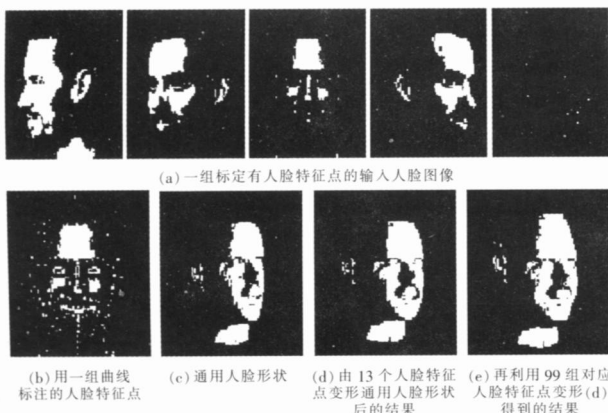


图 2 Pighin 的三维人脸模型建立过程

Akimoto等人^[21]提出基于正侧面两幅人脸图像以及通用人脸模型进行三维人脸建模的方法。其对需要的设备要求不

高,数据采集较少,人工干预较少,能够简单快速地估计人脸的三维形状;但是人脸的纹理恢复不够精细,尤其在正侧面图像中包含的光照条件差异很大的情况下。这给三维纹理的重建带来了很大的困难。

3. 三维人脸模型驱动方法

人脸模型的驱动技术与人脸模型的定义是息息相关的。这些方法有基于几何插值的^[20,22]、基于变形的^[23~25]以及基于参数的人脸动画方法。

基于几何插值的人脸动画方法是人脸动画的一种直觉性的生成方法,通过在关键帧形状之间进行某种形式的加权插值,如线性插值、余弦插值等,生成中间状态的插值图像形状^[20,22]。虽然插值的方法计算复杂度相对较低,动画生成速度快,生成的动画较为逼真,但是生成的表情受到了关键帧的限制,不可能生成关键帧插值范围之外的人脸动作。因此决定了这类方法只适用于少量人脸表情的人脸动画需求。

基于变形的人脸动画方法如 Sederberg等人^[23]提出的 free form deformation (FFD)、Katra等人^[24]提出的 rational free from deformation (RFED),以及 Noh等人^[25]提出的基于径向基函数(radial basis function, RBF)变形的的方法。变形方法能够很好地模拟人脸形状的变化。但由于没有考虑人脸的特殊纹理,不能拟合皱纹、皮肤中的突起以及褶皱,很难生成逼真的人脸动画纹理;并且由于 RFED 是基于曲面变形的,在变形过程中不一定能够保证人脸体积不变。

近年来,研究人员对于基于参数的人脸动画方法作了多方面的研究和尝试。参数化的表达能够通过不同参数值的变化和组合,在相对低的计算复杂度和动画码流率的条件下表达丰富的人脸动作表情。

Parke^[26]对人脸定义了三维的网格模型及其相应的构造参数和表情参数;Blanz与 Vetter^[27]提出了三维可形变模型。在已有的包含 200 个不同人的三维人脸模型(具有三维形状以及纹理的人脸模型)库的基础上定义脸部属性,包括轮廓描述、男子、女子、微笑、皱眉、微胖以及尖勾鼻的差分向量参数表达。三维可形变模型能够在单幅人脸图像的条件下生成逼真的人脸表情图像。但是三维数据以及稠密的光流对应中计算量大、计算复杂度高,并且模型的初始化需要大量的人工干预,因此不适合于需要实时生成的人脸动画应用。

由于预见到说话人脸(talking head)在未来的巨大应用前景,MPEG-4标准在视频编码标准中引进了对合成视觉对象的编码,包括 2D、3D 动画以及人脸面部表情动画^[28~30]。

基于参数化的人脸动画方法能够显式地描述人脸的几何形状以及纹理构成,通过参数的各种不同的值和组合能够产生各种不同的人脸表情动作;并且参数化的控制能够大大减少控制码流,实现在极低比特率的条件下的人脸动画。但是实际的参数化模型很难达到理想参数化水平。目前为止还没有一个很好的提取控制参数的机制,凭主观的判断很难选择到适当的完备参数集合去表达所有可能的人脸表情动作。当两个对相同顶点起冲突作用的参数同时作用于该人脸模型顶点时,很难仲裁如何取舍两个参数或者如何对两个参数加权。因此当参数间有冲突作用存在时,难免产生不自然的表情。如果通过区域划分的方法解决参数冲突问题,又会导致动画时的运动边界问题。

基于图像的人脸动画方法

基于图像的人脸动画方法希望避免三维模型的重构,利用对二维图像某种形式的变形或者拼接处理,直接产生动画的人脸。基于三维模型的方法通过三维形状形变与纹理映射合成动画图像会导致或多或少的人为痕迹。与此不同,基于图像的人脸动画方法生成的动画纹理直接来源于采集的人脸图像,具有很高的纹理真实感。尤其是对于三维模型很难建模的嘴部区域运动纹理,基于图像的方法能够通过采集足够的相应纹理合成逼真的嘴部说话表情纹理。因此在近年来成为人脸动画的一个热门研究方向。这些方法有基于单帧图像形变的、基于关键帧图像的、基于图像拼接的、基于比例图像的以及基于统计模型的。

基于单帧图像形变的人脸动画通过将单帧人脸图像网格化,并根据一定的规则变形网格,拉动人脸纹理曲翘形变而合成动画人脸。Arad等人^[31]提出的基于径向基函数的图像变形由单幅图像实现人脸的表情;Edge等人^[32]基于径向基函数实现了相片般真实的人脸动画系统;Peng等人^[33]实现了语音驱动的人脸动画系统。

基于关键帧的人脸动画方法是计算机动画中最基本并且运用最广泛的方法之一。关键帧的概念来源于传统的卡通片制作。从原理上讲,关键帧插值问题可归结为参数插值问题,传统的插值方法均可应用到关键帧方法中。但关键帧插值又与纯数学的插值不同,它有其特殊性。一个好的关键帧插值方法必须能够产生逼真的运动效果,并能给用户方便有效的控制手段。基于关键帧的人脸动画是Beier和Neely^[34]首先提出来的,其实现的基础是两组拓扑结构相同并一一对应的多边形(三角形)网格模型。Benson^[35]、Gao等人^[36]也基于嘴型图像关键帧的线性组合合成新的动画嘴型;Ezzat和Poggio^[37]基于二维因素图像集合提取相应的关键帧图像,建立了一套二维说话人脸动画系统;2002年,Ezzat等人^[38]建立了一套基于MMM(multidimensional morphable model)的人脸动画系统。

基于图像拼接的人脸动画方法将人脸动画视频的生成问题定义为根据一定的需求和规则来组合和拼接原始的小段分割的人脸图像序列单元,构成新的图像序列。Bregler等人^[39]设计了基于三因素视频重构系统;Cosatto等人^[40]和Graf等人^[41]提出了由基本头部框架、嘴部、双眼和额头等人脸器官组件合成人脸图像的方法。人脸器官组件的不同组合增加了生成动画的多样性;但由于每个器官的运动并非完全独立的,这种主观的分割也有可能带来不自然的动画效果。

基于图像拼接人脸动画方法能够获得高逼真度的人脸动画,但是需要采集大量的图像数据,建立完备的数据集合。在合成时很难实现对动画的调整和控制,并且动画很难推广到其他人脸对象。

基于比例图像的人脸动画方法是一种基于图熵的方法,将两幅已配准的图像相除即得到两者的比例图像。利用比例图像可以在完全基于二维的条件下改变人脸的光照效果^[42,43],也可以在相同光照条件的两幅图像之间捕捉人脸表情动作时纹理的细微变化^[44]。Liu等人^[44]利用表情比例模型能够将一个人脸对象的表情克隆到其他人脸对象,并且能够表现表情的纹理细节,如额头或嘴角由于表情产生的细小褶皱。但是表情

比例模型关注图像纹理的生成,不能很好地拟合动画人脸的形状变化,因此一般不直接用来合成人脸动画图像序列,而是作为人脸动画处理纹理的有力工具,用于增强动画的纹理逼真效果。

基于统计模型的人脸动画方法是通过统计学的方法对训练集图像进行统计分析,构造相应的人脸空间,用人脸空间中的基向量的参数组合(多为线性组合)表示所有人脸图像。在人脸的统计学模型方面做了大量突出工作的是Cootes等人提出的主动形状模型(active shape model,ASM)^[45]、主动表观模型(active appearance model,AAM)^[46]等,能够广泛用于人脸图像的分析、特征点定位、跟踪和合成等领域。ASM与AAM通过对样本集进行主分量分析得到样本的均值与变化模式;然后再用这些提取出来的变化模式线性组合出新样本。这种做法消除了训练样本间的冗余,生成的模型更加紧凑,表示人脸时也更加有效。但是训练ASM与AAM往往需要大量的训练数据,要针对所有的表情动作采集相关训练数据几乎是不可能的。因此它们多用于合成单幅人脸图像,或者应用于人脸特征点的定位。

结束语

现有的人脸动画方法各有特点和相应的适用领域,同时它们在算法的复杂度、获取三维数据设备的易用性或者实现可变姿态的便捷性和鲁棒性等方面还存在不少的问题。不同应用中人脸动画在动画质量(姿态和纹理)、运行性能、建模便捷性以及可操控性等方面提出了不同的需求,因此它们也经常由于难以满足实用的需求而需要相互灵活的结合。

参考文献:

- [1] MEHRABIAN A. Communication without words[J]. *Psychology Today*, 1968, 2(4): 53-56.
- [2] PANDZIC I S, OSTERMANN J, MILLER D. User evaluation: synthetic talking faces for interactive services[J]. *Visual Computer*, 1999, 15(7-8): 330-340.
- [3] PLATT S, BADLER N I. Animating facial expression[J]. *Computer Graphics*, 1981, 15(3): 245-252.
- [4] LEE Y C, TERZOPOULOS D, WATERS K. Constructing physics-based facial model of individuals[C]//Proc of Graphics Interface [S I]: Canadian Information Processing Society, 1993: 1-8.
- [5] WATERS K. A muscle model for animating three-dimensional facial expression[J]. *Computer Graphics*, 1987, 21(4): 17-23.
- [6] MAGNENAT-THALMANN N, PRÉMEAU N E, THALMANN D. Abstract muscle action procedures for human face animation[J]. *Visual Computer*, 1988, 3(5): 290-297.
- [7] TERZOPOULOS D, WATERS K. Analysis and synthesis of facial image sequences using physical and anatomical models[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1993, 15(6): 569-579.
- [8] LEE Y C, TERZOPOULOS D, WATERS K. Realistic modeling for facial animation[C]//Proc of SIGGRAPH '95. 1995: 55-62.
- [9] VIAD M L, YAHYA H. Facial animation with wrinkles[C]//FORSEY D, HEGRON G. Proc of the 3rd Eurographics Workshop on Animation and Simulation. 1992: 11-19.
- [10] KOCH R M, GROSSM H, CARLS F R, et al. Simulating facial sur-

- gery using finite element models [C]//Proc of SIGGRAPH '96. 1996: 421-428.
- [11] LEE S, CHWA K, SHN S. Image metamorphosis using snakes and freeform deformations[C]//Proc of SIGGRAPH '95. 1995: 439-448.
- [12] HONG Peng-yu, WEN Zhen, HUANG T S. iFACE: a 3D synthetic talking face[J]. *International Journal of Image and Graphics*, 2001, 1(1): 19-26.
- [13] JEONG W, KAHLER K, HABER J, *et al*. Automatic generation of subdivision surface head models from point cloud data[C]//Proc of Graphics Interface. 2002: 181-188.
- [14] POSDAMER J, ALTSCHULER M. Surface measurement by space-coded projected beam systems[J]. *Computer Graphics and Image Processing*, 1982, 18(1): 1-17.
- [15] GARCIA E, DUGELAY J L. Low cost 3D face acquisition and modeling[C]//Proc of International Conference on Information Technology: Coding and Computing. 2001: 657-661.
- [16] FUJWARA T, KOSHIMIZU H, FUJIMURA K, *et al*. 3D modeling system of human face and full 3D facial caricaturing[C]//Proc of the 17th International Conference on Virtual Systems and Multimedia. 2001: 625-633.
- [17] GUENTER B, GRIMM C, WOOD D, *et al*. Making faces[C]//Proc of ACM SIGGRAPH Conference on Computer Graphics. 1998: 55-66.
- [18] SHAN Ying, LIU Zi-cheng, ZHANG Zheng-you. Model-based bundle adjustment with application to face modeling[C]//Proc of International Conference on Computer Vision. 2001: 644-651.
- [19] VETTER T, BLANZ V. Generalization to novel views from a single face image[C]//WECHSLER H, PHILLIPS P J, BRUCE V, *et al*. Face recognition: from theory to applications. London: Springer-Verlag, 1998: 310-326.
- [20] PIGHN F. Modeling and animating realistic faces from images[D]. Los Angeles: Institute for Creative Technologies, University of Southern California, 1999: 28-39.
- [21] AKMOTO T, SUENAGA Y, WALLACE R S. Automatic creation of 3D facial models[J]. *IEEE Computer Graphics and Applications*, 1993, 13(3): 16-22.
- [22] WATERS K, LEVERGOOD T M. Deiface: an automatic lip-synchronization algorithm for synthetic faces[R]. Cambridge: Cambridge Research Laboratory, 1993: 31-42.
- [23] SEDERBERG T W, PARRY S R. Free-form deformation of solid geometric models[J]. *Computer Graphics*, 1986, 20(4): 151-160.
- [24] KALRA P, MANGLIA, THALMANN N M, *et al*. Simulation of facial muscle actions based on rational free form deformations[J]. *Eurographics*, 1992, 11(3): 59-69.
- [25] NOH J Y, FIDALEO D, NEUMANN U. Animated deformations with radial basis functions[C]//Proc of ACM Symposium on Virtual Reality and Software Technology. 2000: 166-174.
- [26] PARKE F I. A parametric model for human faces, UTEC-CSc-75-047 [R]. Salt Lake City: University of Utah, 1974: 18-27.
- [27] BLANZ V, VETTER T. A morphable model for the synthesis of 3D faces[C]//Proc of SIGGRAPH. Los Angeles: [s n], 1999: 187-194.
- [28] International Organisation for Standardisation. ISO/IEC 14496-2: 1999/Amd 1: 2000 (E), Information technology coding of audio-visual objects, part 2: visual amendment 1: visual extensions[S]. 2000.
- [29] 姜大龙,王兆其,高文.基于 MPEG-4 的三维人脸动画实现方法[C]//第一届全国虚拟现实与可视化学术会议 (CCVRV '01).北京: [出版者不详], 2001: 49-59.
- [30] 蓝娟,陈益强,徐燕.基于 MPEG-4 和 PCA 的人脸动画合成方法[J]. *计算机工程与应用*, 2004, 40(10): 60-64.
- [31] ARAD N, DYN N, RESFELD D, *et al*. Image warping by radial basis functions: application to facial expressions[J]. *Graphical Models and Image Processing*, 1994, 56(2): 161-172.
- [32] EDGE J D, MADDOCK S. Image based talking heads using radial basis functions[C]//Proc of EGUK2003. 2003: 74-80.
- [33] PERNG W L, WU Yung-kang, OUHYOUNG M. Image talk: a real-time synthetic talking head using one single image with Chinese text-to-speech capability[C]//Proc of the 6th Pacific Conference on Computer Graphics and Applications. Singapore: [s n], 1998: 140-149.
- [34] BEIER T, NEELY S. Feature-based image metamorphosis[C]//Proc of SIGGRAPH '92. 1992: 35-42.
- [35] BENSON P J. Morph transformation of the facial image[J]. *Image and Vision Computing*, 1994, 12(10): 691-696.
- [36] GAO L, MUKAIGAWA Y, OHTA Y. Synthesis of facial images with lip motion from several real views[C]//Proc of Automatic Face and Gesture Recognition. 1998: 181-186.
- [37] EZZAT T, FOGGD T. Mike talk: a talking facial display based on morphing visemes[C]//Proc of Computer Animation Conference. Philadelphia: [s n], 1998: 96-102.
- [38] EZZAT T, GEIGER G, FOGGD T. Trainable videorealistic speech animation[C]//Proc of ACM SIGGRAPH. San Antonio, Texas: [s n], 2002: 388-398.
- [39] BREGLER C, COVELL M, SLANEY M. Video rewrite: driving visual speech with audio[C]//Proc of SIGGRAPH '97. 1997: 353-360.
- [40] COSATTO E, GRAF H P. Sample-based synthesis of photo-realistic talking-heads[C]//Proc of Computer Animation [S 1]: IEEE Computer Society, 1998: 103-110.
- [41] GRAF H P, COSATTO E, EZZAT T. Face analysis for the synthesis of photo-realistic talking heads[C]//Proc of the 4th IEEE Int Conf Automatic Face and Gesture Recognition. Grenoble: IEEE Computer Society, 2000: 189-194.
- [42] SHASHUAN A, RIKL N-RAVIV T. Quotient image: class-based re-rendering and recognition with varying illuminations[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(2): 129-139.
- [43] WEN Zhen, LIU Zi-cheng, HUANG T S. Face relighting with radiance environment maps[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2003: 158-165.
- [44] LIU Zi-cheng, SHAN Ying, ZHANG Zheng-you. Expressive expression mapping with ratio images[C]//Proc of ACM SIGGRAPH Conference on Computer Graphics. Los Angeles: [s n], 2001: 271-276.
- [45] COOTES T F, TAYLOR C J, COOPER D H, *et al*. Active shape models: their training and application[J]. *Computer Vision and Image Understanding*, 1995, 61(1): 38-59.
- [46] COOTES T F, EDWARDS G J, TAYLOR C J. Active appearance models[C]//Proc of the 5th European Conference on Computer Vision. 1998: 484-498.